ChinaXiv合作期刊

基于有线电视互动点播业务的大数据分析推荐系统探析

摘　要：在当今智能化云计算的时代大变革背景下，有线电视也进入了大数据分析的新阶段。互动点播作为有线电视的核心业务之一，将观众从播放控制中解放出来，自由选择喜欢的节目。对这一功能的技术支持就是互动电视传输技术，基于这项技术可以进行大数据的分析等操作。本文围绕有线电视互动点播业务大数据分析系统展开论述，通过推荐数据模型的论述，谈一谈大数据分析推荐系统在有线电视互动点播业务中的应用

关键词：有线电视；互动点播；大数据分析系统
中图分类号：TP391.3　　**文献标识码：**A
文章编号：1671-0134（2017）12-112-02　　**DOI：**10.19483/j.cnki.11-4653/n.2017.01.034

■文 / 王　峰

chinaXiv:202310.03000v1

互动点播业务是广电运营三网融合之后在商业模式上的一个创新之举。在前有移动通讯后有互联网电视的竞争态势下，广电运营的目标是借助互动电视，冲出三网融合之困，与 IPTV、OTT 立于三足鼎立之势。

1. 运用大数据对有线电视互动点播业务进行分析的时代背景

开展互动点播业务，就要开通点播业务功能，这个功能属于付费项目，如何能让用户接受这种付费点播模式呢？从当前市场推广的占有率来看，能够接受付费服务的用户数额较少。首先是常年来我国用户已经习惯了收看原有模式的有线电视节目，而且，对于有线电视的收视习惯也已经固化，忽然让用户自掏腰包购买双向机顶盒，而不使用免费发放的机顶盒，似乎一时无法接受。但是，这个费用与互联网视频的收费相比，是小巫见大巫的。因此收费问题并不是影响有线电视互动点播业务市场占有率低的主要原因。

广电总局关于三网融合的改革进行以来，对于有线电视网络的互动点播业务已经成为了广电运营的主推业务之一。现有的广电商业模式不断进行着创新，互动电视的点播业务正在与 IPTV、OTT 等在市场竞争的残酷态势下争取一席之地。

我国现有的有线电视互动点播业务，真正能够实现付费点播功能还不成熟，能够接受这项业务的用户数量也比较少。这给互动点播业务的拓展、增值业务带来巨大的难度。因此有线电视运营商提供的互动点播业务只是在原有的播出模式上对节目源进行了一定的聚合，距离真正的节目细化、移动电视新业务、节目 EPG 等还相差深远。目前仍然没有摆脱由运营商推送节目，而不是根据用户喜好，对用户群体进行细分后的服务节目的推送模式。导致当前的有线电视业务的使用率盘旋在低位，依然停留在粗放式的经营模式中。

那么从互动点播的节目内容进行分析。有线电视运营商推送的互动点播节目源基本是在原有节目的基础上，再整合了一部分节目，对节目的 EPG 进行了归类和细化，又增加了点播、时移电视等新业务。推送模式依然是传统的，用户的个性、喜好等没有被进行仔细深入的研究[1]。这是导致业务量低的主要原因。也就是说，粗放型的服务方式显然跟不上新时代的需要了。

因此，向互联网视频学习，运用大数据分析系统，将用户喜好、年龄、消费能力等进行挖掘和分析，适时推出适合个人的视频解决方案，是有线电视运营商今后要关注的方向。

2. 大数据分析概述

有线电视的互动点播业务的终端载体为数字机顶盒，通过这一终端，大数据分析系统可以实现的功能包括：

用户机顶盒的 ID、用户的个人情况，如姓名、家庭住址、授权的信息、账户及密码等，都可以通过终端加以采集；媒体的 ID、类型、名称、国家、导演、演员、付费金额、节目简介、高清还是超清等，都可以通过机顶盒终端予以采集；所有的 ID，包括交易、用户、资源，还有点播的时间、点播的量、节目的单价等信息，均可以通过数字机顶盒加以采集；通过 ID 可以对节目的评价、评分等加以采集；对机顶盒遥控器的操作进行数据采集，如播放、暂停等。

通过数字机顶盒等终端采集到的数据，与用户喜好、个性特征、社会阶层等细化的数据采集要求相距甚远。因此还是难以脱离粗放式的数据分析模式。要达到对用户喜好、视频内容详细描述、视频切换频度、节目观看时长等更加精细的数据分析要求，并从中分析用户与数据之间的联系，为有线电视互动点播业务营销提供有利的数据依据，必须要实行数据价值最大化，也就是大数据分析，采用精准推荐模型为数据建模[2]。

2.1 互动点播业务大数据推荐算法

这是为了利用用户网络的分析计算行为，目的就是要推测用户感兴趣的地方到底在哪里。使用到的计算方法包括协同过滤推荐算法、基于内容的推荐算法、基于网络结构的推荐算法等。第一种目前被使用得较为频繁，也就是协同过滤推荐算法：将用户的兴趣模型打造为用户项目关联矩阵，公

ChinaXiv合作期刊

式为：m×n，m 为用户数量，n 是项目的数量。

矩阵的取值与系统的类型设定对应。包括了用户对项目的评分值，用户购买或者浏览项目时对矩阵填充的数据以及根据用户偏好行为形成的目标用户的项目推荐列表。

根据协同过滤推荐算法，我们将有线电视的互动点播业务的数据可以加以整理。以 100 个有线电视互动点播业务用户为计算基数。对不同用户针对同一节目资源的评分数据加以采集 [3]。

对同一资源的收看和评分，用户可以根据用户集 G 进行定义，当有新用户试图点播节目时，采用用户集的方法对新用户的喜好进行预测。

$$r_{c,A} = k\sum_{c|G} sim(c,G) \mid r_{G,A}$$

Sim(c,G) 是采用夹角余弦或者 Pearson 相关的计算。

2. 2 互动点播业务大数据优化算法

有线电视互动点播业务的数据分析目前来说还有很多缺点，如对视频收看的频段、演员的对白兴趣关注度等，还无法掌握到极致，因此，推荐的效果远远达不到令人满意的程度。

这是由于采样设备的局限性造成的结果。机顶盒遥控器对上述问题的解决，只能记录下特定时间段遥控器的操作，如播放、暂停等。对于文档的记录，用户遥控器的数据操作，都还需要进一步从技术层面加大力度，实现数据回传和分析功能，通过推荐的算法，得到更为精准的用户推荐方案 [4]。

我们所要谈的就是这种精准的用户推荐方案——基于内容的推荐算法。这种算法实现了决冷启动，数据稀疏的问题，可以有效补充协同过滤推荐算法。当然，这种算法也有局限性，就是会收到内容提取技术的约束。而且算法复杂，耗费计算能力和时间。

怎么样既不损耗过多的精力，又能达到数据的精准结算呢。本文提出的方案如下：

利用遥控器的操作数据和描述数据的功能，将用户的喜好描述成文档，采用基于内容的推荐算法进行分析，得到数据模型。

从用户喜好描述文档，我们能够看到的数据包括用户喜好度、关键词的出现频率、用户的兴趣点。

再通过协同过滤推荐算法，结合基于内容的推荐算法，对用户的兴趣点进行更加精准的计算，形成更优化的推荐数据模型：将用户喜好度和用户评分显示为文档中不同词汇的权重，然后利用词频差异的特征算法计算用户喜好描述文档，使用向量空间来进行显示，最终完成推荐数据模型。

基于内容的推荐算法可以有效解决冷启动问题、数据稀疏性问题，是基于协同过滤推荐算法的有效补充，但是因为该算法受到内容提取技术的约束，实现较为不易；另外，因为内容提取技术算法较为复杂，需要消耗较多的计算能力和计算时间。为了有效利用既有的有线电视互动点播业务支撑系统，尽量减少新系统的建设成本，一个折衷方案：利用遥控器操作数据和预先制作的视频内容描述数据，按采样时间进行量化分析，形成用户喜好描述文档，再使用基于内容的推荐算法对用户喜好描述文档进行挖掘分析，由此得出推荐数据模型。

3. 有线电视互动点播业务大数据分析优化方案设想

有线电视互动点播大数据来自两个方面，一个方面是通过机顶盒采集，信息内容包括账户和业务数据等。一个方面是对用户操作遥控器采集数据。这类数据是事件类范畴，实时性不高，但是数量庞大。先采用协同过滤推荐算法完成第一次加工，然后采用近似在线的数据处理分析推荐算法完成数据的第二次加工 [5]。

这个解决方案首先是将账户、业务、评价等数据离线存储在平台上，在一定的生产环境下部署有利于性能调整优化的新算法。利用基于协同过滤推荐算法计算框架中的数据并进行加工，结合计算模块中的用户信息进行实时的数据传输与分析，例如初步实时计算基于内容的模块，控制计算量，缩短系统的响应时间，增加推荐精度。

4. 结语

针对有线电视互动点播业务数据的大数据分析，采用基于协同过滤推荐和基于内容的推荐算法是目前较为流行的推荐算法。通过两种算法的混合推荐模型的分析，得到了基于推荐模型的数据算法，为有线电视互动点播的应用提供了大数据技术的支持。媒

参考文献

[1] 章鹏，唐月 . 数据分析技术在有线电视网络经营分析的应用 [J]. 广播与电视技术，2013（3）：113-116.

[2] 苏洪斌，徐鹏飞 . 酒店有线电视与数据网络接入方案探析 [J]. 中国有线电视，2014（6）：767.

[3] 刘海涛 . 有线电视大数据分析平台数据采集 [J]. 广播电视信息，2016（3）：49-50.

[4] 李锐 . 有线电视运营商基于大数据开展社区化的服务 [J]. 电视技术，2014（16）：43-45，60.

[5] 张健 . 基于大数据技术的有线电视推荐系统研究 [J]. 中国有线电视，2016（z1）：362-365.

（作者单位：山东广电网络有限公司日照分公司）

chinaXiv:202310.03000v1